09/402260

The Patent Office

PCT/GB 9 8 / 0 0 9 0 5

0 1 April 1998

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP9 1RH

RECD 30 APR 1998
WIPO PCT

PRIORITY DOCUMENT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

I also certify that the attached copy of the request for grant of a Patent (Form 1/77) bears an amendment, effected by this office, following a request by the applicant and agreed to by the Comptroller-General.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed Andrew Cowie

Dated 2 March 1998

nts Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

02APR97 E264715-2 000056
P01/7700 25.00

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference — MPW/P20180GB

THE PATENT OFFICE D 01 APR 1997 LONDON

2. Patent application number *(The Patent Office will fill in this part)* — 9706529.6 — 01 APR 1997

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

Glaxo Group Limited
Glaxo House
Berkeley Avenue
Greenford
Middlesex UB6 0NN
UK

Patents ADP number *(if you know it)*

If the applicant is a corporate body, give the country/state of its incorporation

4. Title of the invention

Invention

5. Name of your agent *(if you have one)*

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

Kilburn & Strode
30 John Street
London WC1N 2DD
UK

Form 51/77 25/2/97
Andrew J Teuten
Glaxo Wellcome plc
Glaxo Wellcome House
Berkeley Avenue
Greenford
Middx
UB6 0NN

Patents ADP number *(if you know it)* — 125001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	7
Claim(s)	2
Abstract	1
Drawing(s)	3

10. If you are also filing any of the following, state how many against each item.

Priority documents	-
Translations of priority documents	-
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	-
Request for preliminary examination and search *(Patents Form 9/77)*	-
Request for substantive examination *(Patents Form 10/77)*	-
Any other documents *(please specify)*	-

11. I/We request the grant of a patent on the basis of this application.

Signature Gilburn & Strode Date 1/4/97

12. Name and daytime telephone number of person to contact in the United Kingdom

Martin P White - 0171 242 8291

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

INVENTION

The present invention relates to nucleic acid sequencing.

According to one aspect of the present invention there is provided a method for sequencing nucleic acid molecules, the method comprising the steps of:

a) providing at a first location a plurality of single stranded nucleic acid molecules that have the same sequences as one another and that are hybridised to primers in a manner to allow primer extension in the presence of nucleotides and a nucleic acid polymerase;

b) providing at a second location ,which is different from the first location, a plurality of single stranded nucleic acid molecules that have the same sequences as one another, but that may have different sequences from the sequences of the single stranded nucleic acid molecules at the first location, and that are also hybridised to primers in a manner to allow primer extension in the presence of nucleotides and a nucleic acid polymerase;

c) providing each location with a nucleic acid polymerase and a given nucleotide under conditions that allow extension of the primers if a complementary base or if a plurality of such bases is present at the appropriate position in the single stranded nucleic acid molecules;

d) detecting whether or not said nucleotide has been used for primer

extension at each location;

e) repeating steps c) and d) one or more times until a sequence of a desired length (whether full or partial) has been obtained.

Other aspects of the present invention are set forth in the accompanying claims and are described in a non-limiting manner in the following illustrative drawings and example.

Figure 1

Figure 1 illustrates one embodiment of the present invention.

Figure 1(a) shows an example of the annealing of a primer to a template, the primer having a 3' end available for primer extension.

Figure 1(b) shows how labelled bases complementary to the underlined bases shown in Figure 1(a) can be incorporated by a stepwise primer extension cycle using labelled nucleotides.

Figure 1(c) shows how bases with fluorescent labels, which can be incorporated in step (b), can be detected *in situ* by fluorescence measurements.

Figure 2

Figure 2 shows individual spots or "colonies", each comprising large numbers of nucleic acid molecules which are sequenced in the manner illustrated in Figure 2.

Figure 2(a) shows the colonies before fluorescence labelled nucleotides are added.

Figure 2(b) shows the colonies after a fluorescence labelled GTP has been added in the presence of a DNA polymerase and the colonies have been washed to remove any labelled GTP not used in primer extension. The dark colonies are those which have incorporated one or more labelled Gs.

Figures 2(c), (d) and (e) show how the procedure illustrated by Figure 2(b) has been repeated using the fluorescence labelled nucleotides ATP, TTP and CTP respectively (and washing after each cycle to remove nucleotides not used in primer extension).

A fluorescence detection means can be used to distinguish between incorporation of fluorescence labelled A, T, G or C. The latest base to be incorporated in a given colony by primer extension is illustrated in Figures 2(b) to 2(e) by filled circles having different appearances for different bases (G is represented by dark circles, A by circles with oblique shading, T by grid-like shading and C by dots).

Figure 2(f) shows two partial sequences which have been determined by the method illustrated in Figures 2(a) to (e). The sequence shown in lane 1 is GAC, whereas that shown in lane 2 is ATC. For ease of reference, the colonies shown in Figure 2(a) have been identified with numbering corresponding to the lane numbering shown in Figure 2(f). It can thus be seen that two different types of

colony are present in Figure 2(a) (i.e. a given DNA molecule is present in colonies identified with "1" in substantially homogenous form and with a different DNA molecule is present in colonies identified with "2", also substantially homogenous in form).

Of course the present invention can be used to sequence many more than two different nucleic acid molecule sequences.

Figure 3

Figure 3 is referred to in the accompanying example and indicates that the method of the present invention can be successfully applied.

EXAMPLE

This example indicates that successful step-by-step incorporation of nucleotides using the method of the present invention can be achieved. The example uses a sequencing gel for illustrative purposes since such gels are routinely used for sequencing. It will however be appreciated that the present invention is preferably used for *in situ* sequencing of nucleic acid molecules and therefore it does not require the use of sequencing gels.

Experimental procedure :

1. The sequencing cycles were done on single-stranded template molecules bound on magnetic beads.

   The 174 bases DNA fragment was amplified by PCR from the polylinker of plasmid pBlueScript SK-. The forward primer was biotinylated on its 5'-end. The amplified double-stranded DNA was bound on streptavidin coated magnetic beads (Dynabeads M-280, Dynal, Oslo) and denatured with NaOH. After washing, the single-stranded DNA bound to the beads was kept in 10 mM tris pH 8.0.
   The forward primer (5'- GCGCGTAATACGACTCACTA-3') and reverse primer (5'-CGCAATTAACCCTCACTAAA-3') were located on position 621and 794, respectivaly.

2. The reverse primer was used for sequencing. It was labeled on its 5' end with radioactive phosphate, and annealed to the single-stranded DNA template. The primer molecule is expected to be extended with the following bases : GGGAACAAAAGCTGGAG...

   The primer end-labeling was done with kinase enzyme (Pharmacia) in presence of $^{32}P\ \gamma$ dATP. After purification on Sephadex G-25 spin column, the primer were conserved at -20°C.

The annealing was performed by heating the mixture of primer and template molecules for 5 min at 70°C and slow cooling for 2 hours.

3. The sequencing reaction were performed by successive cycles of primer extension (in presence of only one type of nucleotide), aliquot removal for analysis, and washing. Each cycle was done using a different nucleotide in solution, using the arbitrary order of A, T, G, C, A, T, G, C, A, T, etc..

   We choose the T7 Sequenase II (Amersham) as the enzyme for polymerization reactions. In contrast to recommendations of the manufacturer the reaction mix only contained 0.1 u/µl Sequenase and 100 nM of a single type of normal deoxynucleotide. Four different mixes were prepared, each with a different type of nucleotide.

   A cycle started with the removal of buffer from the template-bound beads using a magnet. Ice-cold reaction mixture was added. After 5 seconds an aliquot was collected and kept on ice in 10 volumes of stop buffer containing formamide and EDTA. Ten volumes of washing buffer (10 mM Tris pH 8.0) was added to the remaining reaction mixture and immediately removed after magnetic separation of the template-bound beads. Washing was repeated and next cycle was initiated.

   At each cycle, a smaller volume of fresh reaction mixture was used, according to the remaining amount of template after aliquot removal and washing losses.

4. The aliquots collected after each cycle were then analyzed on denaturing sequencing gel. After migration, the gel was autoradiographed to reveal to position of each band.

   The aliquots in stop buffer were denatured for 5 min at 95°C and immediately cooled on ice before loading on 0.75 x 180 x 320 mm acrylamide gels containing 1xTBE buffer and 7 M urea. The system was model SE420 from Hoefer Scientific Instruments. Migration was at 2000 V for 5 min and 250 V overnight.

   As a control, unreacted oligonucleotides were applied on gel (lane 1), as well as the product of a reaction performed in presence of all 4 nucleotides, to give rise to fully extended molecules (lane 2).

   After migration, the gel was fixed in 10% glacial acetic acid and 10% methanol for 30 min. Still humid, it was transferred in a plastic bag in autoradiography cassette to expose a Kodak X-OMAT AR film for 9h at 25°C.

Results and discussion:

The very high sensitivity of radio-active label, associated with gel analysis of extended fragments after each cycle allowed careful monitoring of the reaction. As shown on figure 3it was demonstrated that:

- Conditions were identified which permit correct polymerization in presence of only one type of nucleotide. The polymerization proceeds completely to the last base to be incorporated (lane 5), but is blocked for (further) extension if the complementary base is not present on template molecule (lane 3, 4, or 5)
- There is no misincorporation, as seen on lanes 3, 4, or 6.
- There is no 3'-5' exonuclease activity (*e.g.* fragments of shorter size are not detected on lane 3 or 4).

- Successive cycles can be carried out correctly

  Only fragments of the correct size are observed. There are three exceptions : (1) a small proportion of the primer molecules were not extended during first cycle, probably because of 3' end damage of the oligonucleotide. (2) A proportion of full length molecules are observed after cycle 13, which suggests that washing was not complete and that the 4 nucleotides were present in solution. This problem is not considered as relevant, because it was not observed between cycles 4 and 12, even though the 4 different nucleotides had already been used for reactions. The washing steps should also be facilitated when the experiment will be repeated with template molecules bound on plastic surface. (3) Faint bands of smaller size than expected could be detected after longer exposure, which suggests uncompleted reaction. However, the proportion is so low that it should not interfere with *in situ* detection.

  Only 19 cycles were carried out in this preliminary experiment, and we see no reason why more cycles (up to 50 or up to 100 or more) could not be performed successfully.

This experiment demonstrated that no only re-sequencing (when only the expected nucleotide is used for polymerization), but that also *de novo* sequencing can be carried out efficiently with this method.

After 20 cycles of *de novo* sequencing of an unknown template molecule, the minimal number of bases which could be sequenced is of 5, the maximal number depending on the base order and on the presence of runs of the same base. In this experiment, as much as 17 bases were read in 19 cycles.

It should also be mentioned that the template molecule used here is considered as difficult to sequence because of its high content of palindromic regions, which give rise to secondary structures in single-stranded DNA. This potential difficulty appeared not to be a problem in our experiment.

Conclusion:

We demonstrated here that sequencing can be performed stepwise using a single type of normal nucleotide.

The method described here permits either re-sequencing or *de novo* sequencing. It brings a new tool for simultaneous mutation detection and identification on a very large scale and on various genetic locations.

Claims

1. A method for sequencing nucleic acid molecules, the method comprising the steps of:

a) providing at a first location a plurality of single stranded nucleic acid molecules that have the same sequences as one another and that are hybridised to primers in a manner to allow primer extension in the presence of nucleotides and a nucleic acid polymerase;
b) providing at a second location, which is different from the first location, a plurality of single stranded nucleic acid molecules that have the same sequences as one another, but that may have different sequences from the sequences of the single stranded nucleic acid molecules at the first location, and that are also hybridised to primers in a manner to allow primer extension in the presence of nucleotides and a nucleic acid polymerase;
c) providing each location with a nucleic acid polymerase and a given nucleotide under conditions that allow extension of the primers if a complementary base or if a plurality of such bases is present at the appropriate position in the single stranded nucleic acid molecules;
d) detecting whether or not said nucleotide has been used for primer extension at each location;
e) repeating steps c) and d) one or more times until a sequence of a desired length (whether full or partial) has been obtained.

2. A method according to claim 1, wherein all or part of the sequence that is obtained in step e) is converted to provide a complementary sequence thereto.

3. A method according to any preceding claim, wherein if the given nucleotide has been used in primer extension in step d) then this step includes the step of detecting how many of the given nucleotides have been used per extended primer.

4. A method according to any preceding claim, wherein after step c) excess nucleotides that have not been used in primer extension are removed (e.g. by washing).

5. A method according to any preceding claim, wherein the nucleotides are labelled with a detectable label and detection of the label is used in step d)

6. A method according to any preceding claim, wherein step d) uses absorption or emission spectrometry

7. A method according to any preceding claim, wherein said single stranded nucleic acid molecules, said primers or both of the aforesaid are immobilised.

8. A method according to any preceding claim that is used to fully or partially sequence 10 or more nucleic acid molecules simultaneously

9. A method according to any preceding claim that is used to fully or partially sequence 100 or more nucleic acid molecules simultaneously

10. A method according to any preceding claim that is used to fully or partially sequence 1000 or more nucleic acid molecules simultaneously

11. A method according to any preceding claim, wherein each of four different nucleotides is used in primer extension.

12. A method according to claim 11, wherein said four different nucleotides are used in a predetermined order in repeated cycles.

13. A method according to claim 11 or claim 12, wherein the nucleotides are ATP, TTP, GTP, CTP, or labelled derivatives thereof

14. A method according to claim 11 or claim 12, wherein the nucleotides are ATP, UTP, GTP, CTP, or labelled derivatives thereof.

15. A method according to any preceding claim, wherein the detection step is carried out without moving the nucleic acid molecules from the different locations.

16. An apparatus for performing a method according to any preceding claim, the apparatus comprising a plurality of nucleotides, a nucleic acid polymerase and detection means for performing step d) of claim 1, the detection means being adapted to distinguish between said different locations.

17. An apparatus according to claim 16 comprising means for removing excess nucleotides from the different locations (e.g. washing means).

18. An apparatus according to claim 16 or claim 17 that is automated to allow repeated cycles of primer extension and detection.

19. The invention substantially as hereinbefore described.

ABSTRACT

INVENTION

Nucleic acid molecules at distinct locations can be sequenced simultaneously. Primers that are annealed to the nucleic acid molecules are provided. Each location is then provided with a nucleic acid polymerase and a nucleotide. It is then determined whether or not the nucleotide has been used in primer extension.

FIGURE 1

n situ Sequencing

```
Primer :          5'-gactagcgtcat-3'
                     IIIIIIIIIII
Template :  3'-ggatgctgatcgcactattgatgggcacgaactca-5'
```

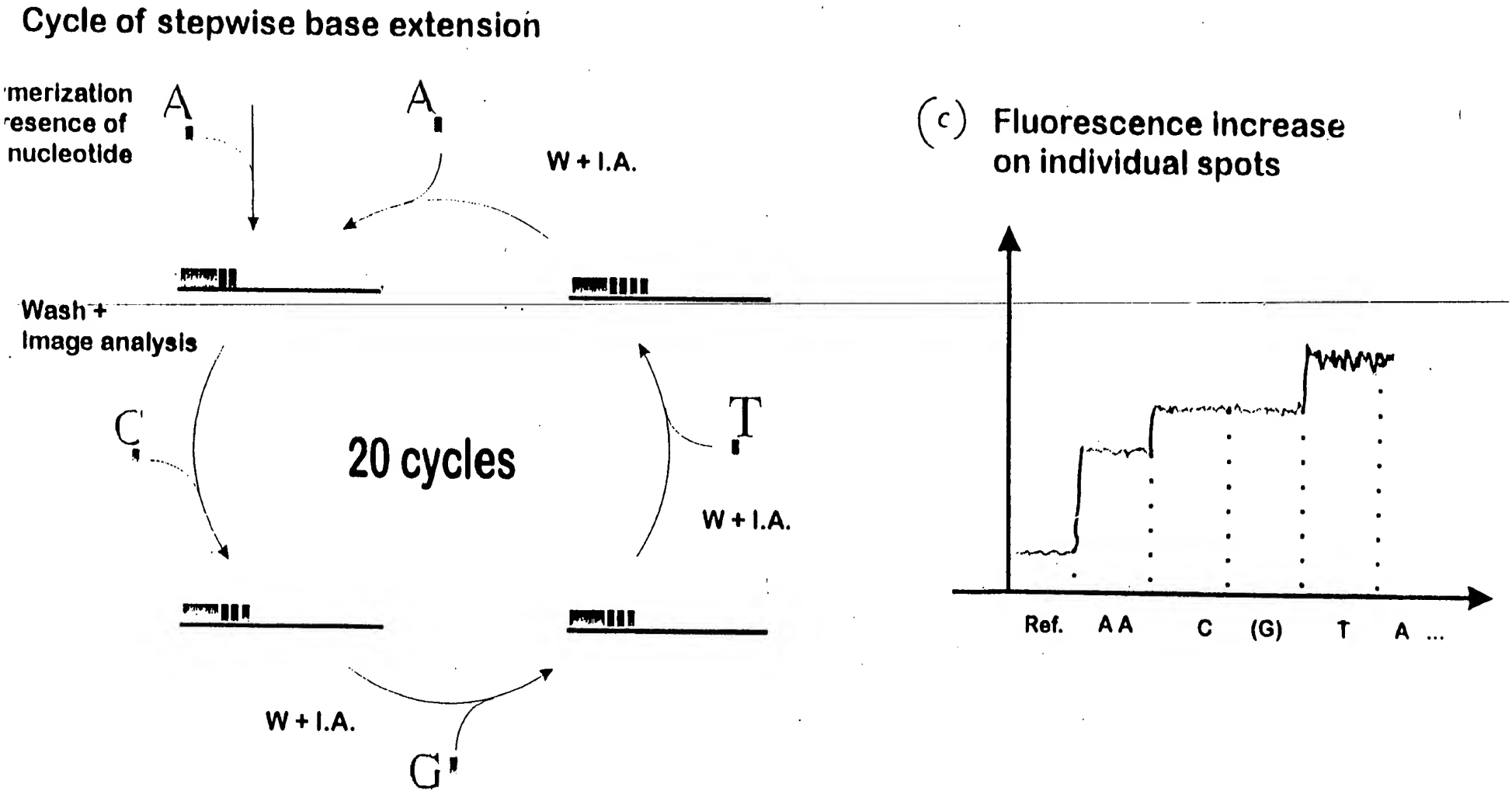
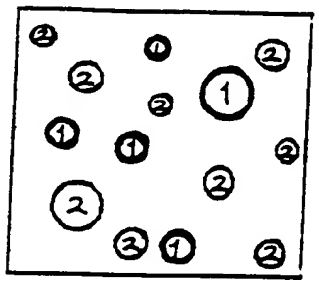


FIGURE 2

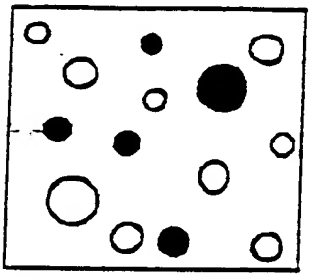
in situ Sequencing



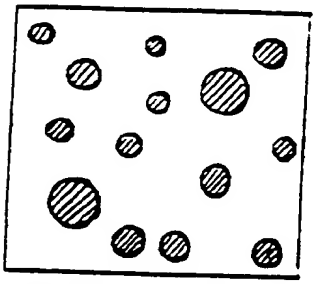
Colony types:
1 or 2

a)

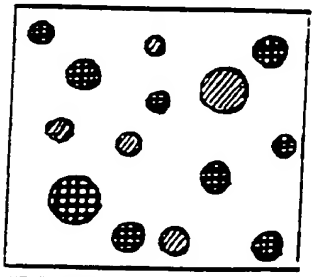
PCR colonies



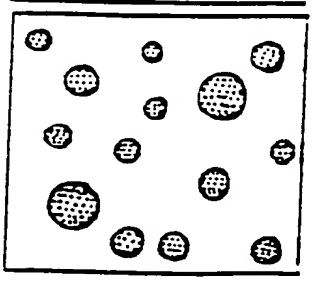
b) G



c) A



d) T



e) C

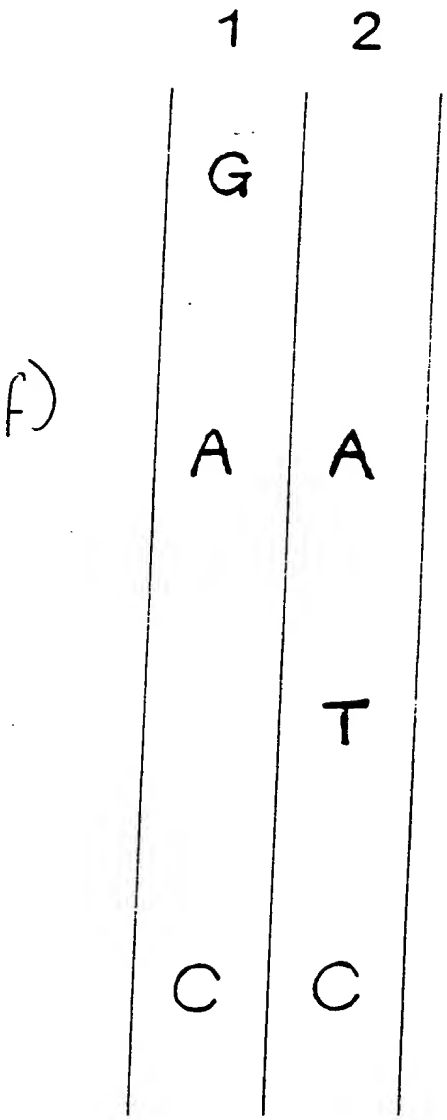
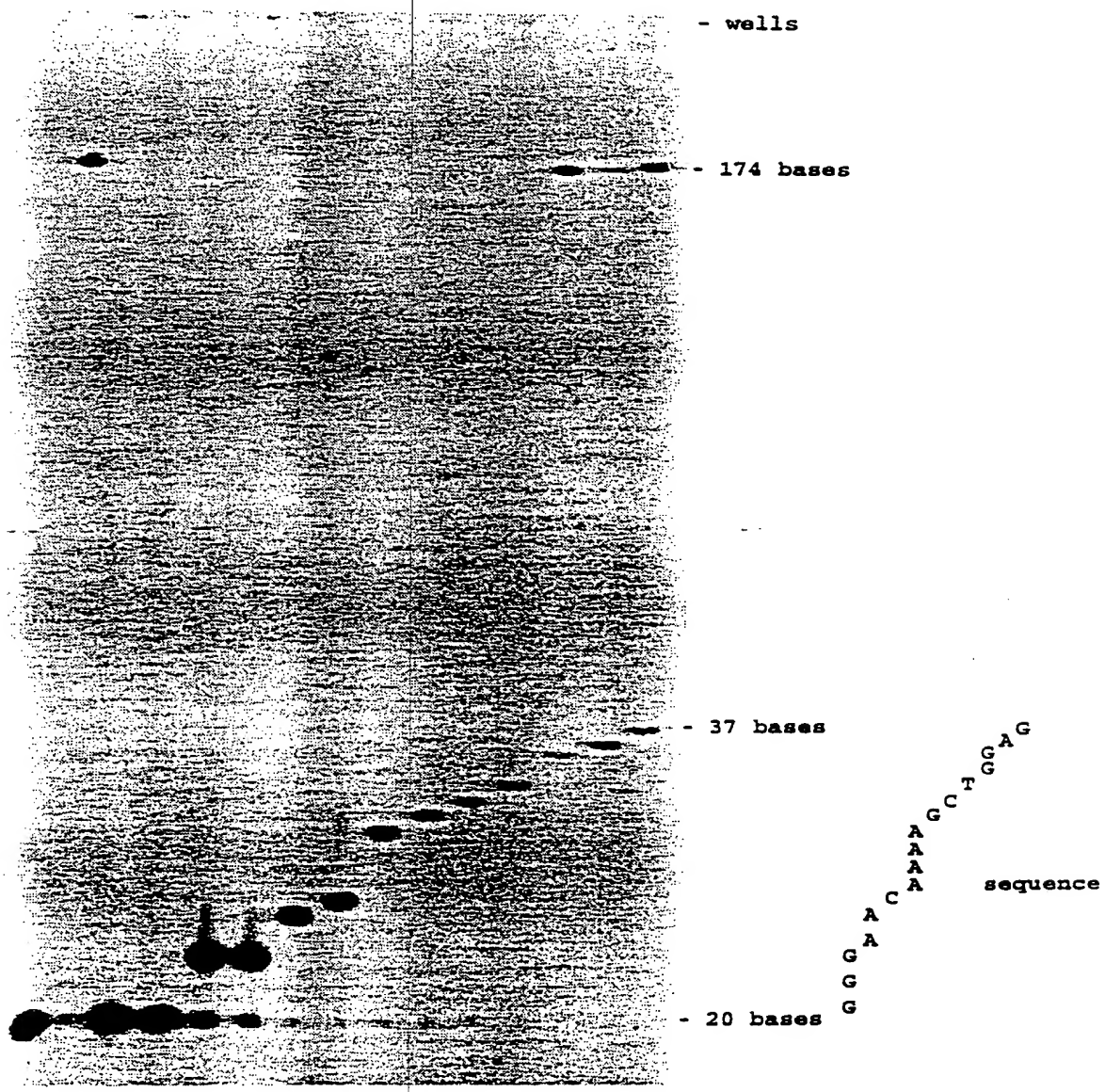


FIGURE 3

in situ Sequencing : *de novo* Sequencing



Lane	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Nucleotide added	-	N	A	T	G	C	A	T G C	A	T G	C	A T	G	C A	T G
Step		1	1	2	3	4	5	6 7 8	9	10 11	12	13 14	15	16 17	18 19

0.75 mm 15% Acrylamide gel, TBE + 7M Urea